Informatics in Medicine Unlocked 10 (2018) 45–49

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Trends and seasonality extracting from Home Blood Pressure Monitoring readings

G.P. Chuiko [a], O.V. Dvornik [a,*], I.A. Shyian [a], Ye.A. Baganov [b,c]

[a] Petro Mohyla Black Sea National University, 68 Desantnikov St., 10, Mykolayiv, 54003, Ukraine
[b] Kherson National Technical University, 24 Berislavske Shose, Kherson, 73008, Ukraine
[c] Kherson State Maritime Academy, 20 Ushakova Avenue, Kherson, 73000, Ukraine

A R T I C L E I N F O

Keywords:
Home blood pressure monitoring
Singular spectrum analysis
Trend
Long-term variability
Computer signal processing

A B S T R A C T

Purpose: The aim is the first application of Singular Spectrum Analysis for computer processing of Home Blood Pressure Monitoring (HBPM). An illustration of the method advantages is the additional objective.
Method: The Singular Spectrum Analysis (SSA) is a way of trend and seasonality extraction. SSA is suitable for short series with noises. It is also useful for the variability analysis. As well, it is simply programmable in a high-level programming language. Here Maple 18 was in use.
Results: Trends and the slowest oscillations were obtained within a case study. The error of SSA estimations turned out to be close to the exactness of measuring. However, these errors are in good agreement with the known descriptors of the short-time variability. The periods of the slowest oscillations were different for the systolic and the diastolic pressures. Seasonality was about six months and about two months respectively. We found the amplitude of the slowest oscillations of heart rate varies in time. This fact is in contrast to the blood pressure fluctuations, for which the amplitudes are stable.
Conclusions: SSA is a promising tool for HBPM data processing. It ensures smooth and readable trends as well as shows the long-term variability of series. This information can be usable for clinical decision-making and prognostics.

1. Introduction

Home Blood Pressure Monitoring (HBPM) is a known healthcare tool [1–3]. The aim of HBPM is the prophylaxis of hypertension that is a global trouble. Monitoring is a way to know how the pressure and heart rate are changing during a long period. It allows tracing how they react to lifestyle or medical treatment.

The advantages and conditions of HBPM were studied in detail in Ref. [3]. Monitoring requires the correct processing of data at first. The reliable information about trends is a key for clinical decision-making and prognostics. The long-term variability of pressure and heart rate are also in use. Besides, it is desirable getting these results in the computer graphs form.

The common attention never paid enough to computer processing of HBPM. We can refer only a few papers [4–7], which are more or less involved in this theme. The appropriate ways of the extraction of trends and seasonality from HBPM still have been studied.

One of the promising methods for trend extracting is Singular Spectrum Analysis (SSA) [8–11]. SSA decomposes a series into the sum of few parts, which are usually interpreted as the “trend”, “seasonality”, and residuals. These residuals are often termed as “noise” [9,10].

The SSA is applicable for short series with noises like HBPM readings. It means a series with length starting with a few dozen of samples. SSA does not demand any assumptions about signal or noise parameters [9]. A user does not need to know a parametric model of the considered time series [10]. This is the extra advantage of the method because such info is unreal or unreliable for HBPM.

The algorithms for SSA realization, suitable for computing, have been described in Ref. [11]. These algorithms need powerful program package of the linear algebra. Due to this reason as well as due to excellent graphics the authors used Maple 18 [12] to realize SSA algorithms.

The aim of this paper is the applicability of SSA for computer processing of HBPM. The extracting of trends and slowest oscillations (seasonality) were performed using the cited above HBPM data [7].

* Corresponding author.
E-mail addresses: genchuiko@gmail.com (G.P. Chuiko), olga.dvornik@chmnu.edu.ua (O.V. Dvornik), iadanishevskaya@gmail.com (I.A. Shyian), ye.baganov@gmail.com (Ye.A. Baganov).

https://doi.org/10.1016/j.imu.2017.12.001
Received 28 September 2017; Received in revised form 19 November 2017; Accepted 2 December 2017
Available online 7 December 2017

2. Measurements, methods and basic algorithm

The patient is the 67-year-old male with the long history of atherosclerosis. Nebival (5 mg) and Lozap (50 mg) have been daily drugs.

The HBPM was done on semi-automatic tonometer Microlife BP 3AG1. The data sets were the systolic (SBP) and diastolic (DBP) blood pressures as well as the heart rate (HR). The guaranteed accuracy of the tonometer is ± 3 mmHg for the blood pressure and $\pm 5\%$ for the heart rate. The heart rate was measured in "beats per minute" (bpm).

The monitoring was organized according to the protocol [1]. The patient was properly positioned, the cuffs were on the naked arm, and the measurements were conducted after the bathroom [2]. Three trials with interval about 1 min were performed each time. The mean of them was taken as the result. The used device allows checking the pulse irregularities automatically. A few such cases had happened and trials were repeated.

The two-day break was set between two serial trials. HBPM performed during over a year (more exact during 384 days). So, each of three data sets consisted of $N = 128$ samples. These data sets have been presented earlier in Ref. [7].

The category "short-term" means below the interval of few days or up to one week. Meanwhile "long-term" notes here a few weeks duration or longer. The equal terms are "high-frequency" and "low-frequency".

Maple 18 has power program packages for linear algebra and statistical calculations [12]. These packages were in use for computer processing of data sets.

The core algorithm of SSA decomposition consists of four key steps [9–11]:

1 Embedding;
2 Singular Value Decomposition;
3 Grouping procedure;
4 Hankelization (anti-diagonal averaging procedure).

The main purpose of SSA is to decompose the original series into a sum of series, so that each component in this sum can be identified as a trend or periodic (or quasi-periodic) component (e.g., amplitude-modulated), or residuals alias noise [10].

We are going to realize all mentioned above as applied to the HBPM data.

3. Realization of SSA algorithm

3.1. Embedding and trajectory matrix

Let $F = (f_1, \ldots, f_N)$ is a real-valued time series with the length of $N > 2$. Let a window length is L $(1 < L < N)$. The embedding converts the original series to the set of lagged vectors. This procedure creates the sequence of $K = N - L + 1$ lagged column-vectors $\mathbf{x}_i \in \mathbf{R}^L$:

$$\mathbf{x}_i = (f_1, \ldots, f_{i+L-1})^T; \quad 1 \le i \le K \tag{1}$$

The trajectory matrix $\mathbf{X}$ of the series F takes in the vectors (1) as columns:

$$\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_K) = \begin{pmatrix} f_1 & f_2 & \ldots & f_{K-1} & f_K \\ f_2 & f_3 & \ldots & f_K & f_{K+1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ f_L & f_{L+1} & \ldots & f_{N-1} & f_N \end{pmatrix} \tag{2}$$

Note, the trajectory matrix has equal values along the anti-diagonals $2 \le i + j = const \le N + 1$. So, it is a Hankel matrix [8–11]. Samples of the series $F = (f_1, \ldots, f_N)$ define this matrix. The matrix dimensions are (L, K) and depend on the choice of the window length.

The window length is the important parameter of the embedding step [9]. The window length L should be large enough. Under this condition, each of column-vectors (1) presents the significant part of the series.

The trajectory matrix (2) possesses a special symmetry property. Indeed, its transposed matrix $\mathbf{X}^T$ is the trajectory matrix of the same series F. However, the transposed matrix has window length equal to K and dimensions (K, L). Obviously, if accept $L < K$ then $L \le \frac{N}{2}$. Thereby, the choice of a larger window length is inexpedient.

The large length of window is desirable if the extraction of the series trend is the aim. The larger is the window length the more detailed will be the decomposition of the series. Thus, the choice of maximal possible length of window provides the most detailed decomposition [8–11]. That is why the length of the window was chosen equal to $L_{\max} = \frac{N}{2} = 64$. Then $K = \frac{N}{2} + 1 = 65$.

3.2. Singular Value Decomposition

The Singular Value Decomposition (SVD) is a process of factorization of a matrix. SVD lets the next representation of the trajectory matrix:

$$\mathbf{X} = \mathbf{USV} \tag{3}$$

Here $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_L)$ is the (L, L) orthogonal matrix that columns are the left singular vectors. $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_K)^T$ is the (K, K) orthogonal matrix that rows are the right singular vectors. $\mathbf{S}$ is the (L, K) diagonal matrix. Its non-negative real diagonal elements $S_{ii} = s_i$ $(1 \le i \le L)$ are the singular values of the matrix $\mathbf{X}$. These singular values are the square roots of the eigenvalues of covariance matrix $\mathbf{C} = \mathbf{XX}^T$. All eigenvalues are real and non-negative. Hence, their square roots are also real and non-negative. The singular values are usually presented as a sorted in descending order list: $s_1 \ge s_2 \ge \ldots \ge s_L$.

A finite sum of one-rank (L, K) matrices presents the trajectory matrix [8–10]:

$$\mathbf{X} = \mathbf{X}_1 + \ldots + \mathbf{X}_L = \sum_{i=1}^{L} \mathbf{X}_i \tag{4}$$

The Cartesian product of the left and the right singular vectors defines a one-rank matrix $\mathbf{X}_i$:

$$\mathbf{X}_i = s_i \left(\mathbf{u}_i \otimes \mathbf{v}_i^T \right); \quad 1 \le i \le L \tag{5}$$

Here $\mathbf{u}_i$ is the column-vector of $\mathbf{U}$ and $\mathbf{v}_i$ is the corresponding row-vector of $\mathbf{V}$. The singular value s_i belongs to the above-mentioned descending ordered list.

The matrices (5) have rank 1, so they are called elementary matrices. The triples $\{s_i, \mathbf{u}_i, \mathbf{v}_i^T\}$ with fixed index are called singular triples.

3.3. Grouping procedure

One can cut off "the tail" of series (4). Of course, the truncated series only estimates the trajectory matrix. The problem is in the decision how many matrices of series (4) one has to keep. What part of information will be kept and what one will be lost after truncation?

The goal of the grouping of the singular triples is the proper separation of the initial series F into several ones. A singular triple $\{s_i, \mathbf{u}_i, \mathbf{v}_i^T\}$ fully defines each matrix $\mathbf{X}_i$. Thus, one can talk about the grouping of the triples instead of the elementary matrices [9–11].

One simple and commonly accepted [8–10] ratio estimates the relative contribution of a singular triple in the sum (4):

$$\alpha_i = \frac{s_i^2}{\sum_{i=1}^{L} s_i^2} \tag{6}$$

A few the first triples often are enough to obtain the well-detailed trends [8–11].

There the question arises: what does it mean "a few"? The answer was obtained a long time ago [13,14]. The neglecting of all triples with singular values less than the mean singular value is the advice of so-called